



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### **Technical note: A note on the differential impact of wrong and missing sire information on reliability and gain**

**Citation for published version:**

Woolliams, J 2006, 'Technical note: A note on the differential impact of wrong and missing sire information on reliability and gain', *Journal of Dairy Science*, vol. 89, no. 12, pp. 4901-4902.

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Journal of Dairy Science

**Publisher Rights Statement:**

© American Dairy Science Association, 2006

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## Technical Note: A Note on the Differential Impact of Wrong and Missing Sire Information on Reliability and Gain

J. A. Woolliams

Roslin Institute (Edinburgh), Roslin, Midlothian, United Kingdom

### ABSTRACT

This note analytically derives the impact that wrong and missing sire information (WSI and MSI, respectively) has on the reliability of predicting merit and gain compared with perfect information. In particular, for small WSI and MSI, WSI was shown to have twice the impact of MSI for both reliability and gain, and the impact of both WSI and MSI increased as the reliability of predicting merit with perfect information decreased. The overall impact on the efficiency of gain for small WSI and MSI was half the overall impact on reliability. **Key words:** reliability, gain, wrong sire information, missing sire information

A recent article by Sanders et al. (2006) highlighted the distinction between the fraction of a sire's total progeny with wrong sire information (**WSI**) and the fraction with missing sire information (**MSI**) when considering the impact of imperfect information on genetic evaluation and gain. They showed, primarily by simulation, a difference between these situations and indicated that WSI was more deleterious than MSI. An Appendix to the paper quantified the impacts directly, but the authors introduced an error in the results for genetic gain. Therefore, this short note reviews this Appendix and its conclusions.

Define  $R$  to be the reliability that would be obtained without pedigree errors, with  $R_e$  the reliability having introduced a number of pedigree errors attributable to WSI or MSI. If  $N$  is the number of potential progenies per sire and  $t$  is the intraclass correlation, then

$$R = N(N + \lambda)^{-1}$$

where

$$\lambda = (1 - t)t^{-1}$$

(e.g., Mrode, 1996). Sanders et al. (2006) develop

$$R_e = (1 - \text{MSI})N / [(1 - \text{MSI})N + \lambda_e]$$

where

$$\lambda_e = [1 - (1 - \text{WSI})t] / [(1 - \text{WSI})^2t].$$

This form is an extension of previous work summarized by Visscher et al. (2002): the MSI affects the reliability of bulls through the amount of data presented for evaluation, whereas the WSI affects reliability through the parameters used in the evaluation. The parameters are assumed to have been estimated previously from a different data set with the same degree of WSI.

Let  $E_R = R_e/R$ . Then defining  $x = (1 - \text{MSI})$  and  $y = (1 - \text{WSI})$ ,

$$E_R = xy^2[t(N - 1) + 1] / [yt(xyN - 1) + 1].$$

Differentiating with respect to  $x$  gives

$$dE_R/dx = E_R/x - E_R y^2 t N / [yt(xyN - 1) + 1].$$

The differential of  $E_R$  with respect to MSI, following the chain rule for differentiation, is

$$[dx/d(\text{MSI})][dE_R/dx] = -dE_R/dx.$$

With full information, that is,  $x = y = 1$ ,  $E_R = 1$ , then

$$dE_R/dx = 1 - tN / [t(N - 1) + 1] = 1 - R,$$

and

$$dE_R/d(\text{MSI}) = -(1 - R).$$

Differentiating with respect to  $y$  gives

$$dE_R/dy = 2E_R/y - E_R t(2xyN - 1) / [yt(xyN - 1) + 1].$$

The differential of  $E_R$  with respect to WSI, again using the chain rule, is  $-dE_R/dy$ . With full information, that is,  $x = y = 1$ ,  $E_R = 1$ , then

$$dE_R/dy = 2 - 2tN[1 - (2N)^{-1}] / [t(N - 1) + 1] = 2\{1 - R[1 - (2N)^{-1}]\}.$$

Received May 1, 2006.

Accepted August 1, 2006.

E-mail: john.woolliams@bbsrc.ac.uk

Therefore, with moderate to large  $N$ ,

$$dE_R/d(WSI) \approx -2(1 - R).$$

This approximation remains good for  $N$  as small as 5, because the exact derivative is then  $-2(1 - 0.9R)$ . For  $N = 1$ ,

$$dE_R/d(WSI) = -(1 - R) - 1,$$

so even at the extreme, the magnitude of  $dE_R/d(WSI)$  is greater than that of  $dE_R/d(MSI)$ .

The following conclusions may be drawn for the reliability of predicting the merit of a sire: 1) For small MSI and WSI, the impact of WSI is approximately twice as large as that of MSI in reducing reliability; and 2) the impact of WSI and MSI on  $E_R$  increases as  $R$  decreases, so that relatively small  $N$  and small  $t$  (i.e., small  $h^2$ ) will lead to greater reductions in  $E_R$ .

Define  $E_G$  to be the relative efficiency of gain, defined as the ratio of gain achieved with MSI and WSI to gain achieved with perfect information. The impact of MSI and WSI on  $E_G$  was incorrectly developed by Sanders et al. (2006). Because  $\Delta G \sim R^{1/2}$ , the impact on the efficiency of a breeding plan may be inferred from  $E_G = E_R^{1/2}$  (Visscher et al., 2002). By using the chain rule for differentiation,

$$dE_G/dx = \frac{1}{2}E_R^{-1/2} dE_R/dx,$$

and analogously for  $y$ . Therefore, the relative sensitivity of  $E_G$  to MSI and to WSI is determined by the relative sensitivity of  $E_R$  to MSI and to WSI, and remains unchanged. Thus, for small MSI and WSI, WSI has approximately twice the impact of MSI, not 1.4-fold, as indicated by Sanders et al. (2006). Similarly, the po-

tentiating factors of relatively small  $N$  and small  $t$ , resulting in small  $R$ , will also result in a greater impact of MSI and WSI on gain. However, the reduction in  $E_G$  will be approximately one-half the reduction observed in  $E_R$  for small MSI and WSI, because  $E_R$ , and hence  $E_R^{-1/2}$ , is close to 1. For example, a change in MSI (or WSI) sufficient to reduce  $E_R$  by 0.01 will reduce  $E_G$  by only 0.005.

These findings provide some guidance on managing the risks of imperfect pedigree information. In the present situation, the impact is measured by the derivatives derived above, giving a value to the loss of gain for an increment of error. The full form of the derivatives can be used for all values of  $N$ ,  $t$ ,  $x$ , and  $y$ , but the particular results for  $x$  and  $y$  close to 1 have been given because their form provides some clear insight into the problem and they represent a limiting form on the approach to perfect information. Risk is defined by both the impact of an error and how likely the error is to occur, and the results suggest that where subjective assessments of pedigree are unable to identify a sire as being more likely than not, then a missing value has less risk than using the most likely among several potential sires. This situation might be avoided where evaluations fully account for degrees of uncertainty in pedigree, such as might be achieved through the use of Monte Carlo Markov chain methods.

## REFERENCES

- Mrode, R. A. 1996. *Linear Models for the Prediction of Animal Breeding Values*. CAB International, Wallingford, Oxon. UK.
- Sanders, K., J. Bennewitz, and E. Kalm. 2006. Wrong and missing sire information affects genetic gain in the Angeln dairy cattle population. *J. Dairy Sci.* 89:315–321.
- Visscher, P. M., J. A. Woolliams, D. Smith, and J. L. Williams. 2002. Estimation of pedigree errors in the UK dairy population using microsatellite markers and the impact on selection. *J. Dairy Sci.* 85:2368–2375.